

**Analisis Kualitas Tes Buatan Guru Mata Pelajaran Matematika Kelas V SD Negeri  
Di Kecamatan Wotu Kabupaten Luwu Timur  
Tahun Pelajaran 2018/2019**

Fitriah

Program PascaSarjana, Universitas Negeri Makassar, Makassar  
fitriah.tosanabe@gmail.com

**Abstrak :** Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal ulangan semester ganjil yang ditinjau dari validitas, reliabilitas, tingkat kesukaran, daya beda, efektif tidaknya pengecoh soal buatan guru Mata Pelajaran Matematika tingkat SD Kabupaten Luwu Timur. Penelitian ini bersifat *ex-post facto*. Data penelitian ini berupa: tahapan pembuatan soal, kisi-kisi soal semester, kartu soal, soal semester dan lembaran hasil jawaban siswa. Teknik pengumpulan data menggunakan dokumentasi. Teknik analisis data yang mengungkap kualitas soal dilakukan penelaan oleh pakar, dan analisis secara empirik berdasarkan teori tes klasik. Hasil penelitian menunjukkan bahwa : (1) Kualitas butir berdasarkan validitas isi dengan koefisien Gregory =  $0,87 > 0,75$ , maka disimpulkan bahwa hasil evaluasi antar pakar memiliki relevansi yang kuat; (2) berdasarkan validitas kriteria, butir yang valid 56,7% (17 butir); (3) kualitas butir berdasarkan reliabilitas, memiliki koefisien reliabilitas yang rendah yaitu 0.667%; (4) Kualitas soal berdasarkan daya pembeda, butir yang memiliki daya pembeda buruk/jelek = 2 butir atau 6,7%, kurang baik = 4 butir atau 13,3%, baik = 7 butir atau 23,3%, dan sangat baik = 17 butir atau 56,7%; (5) Kualitas butir berdasarkan tingkat kesukaran, 56,7% (17) butir yang tergolong sukar, dan 43,3% (13) butir tergolong sedang; (6) Kualitas soal berdasarkan analisis efektivitas pengecoh, terdapat 4 butir yang memiliki pengecoh tidak berfungsi.

**Kata Kunci :** Analisis Kualitas Soal Semester, dan Teori Klasik

**Abstract:** The study aims at examining the quality of test items of odd semester based on validity, reliability, level of difficulty, the effectiveness of distracter of test made by Mathematics subject teachers in elementary level in East Luwu district based on classical theory. The study is *ex-post facto*. Data of the study the stages of test made, question grid, test card, semester test, and answer sheet. Data collection technique employed documentation. Data analysis technique used to discover the quality of test; theoretically, it was reviewed by the experts, and empirically it was based on classical test theory. The result of the study reveal that (1) the quality of test items based on content validity with Gregory coefficient =  $0.87 > 0.75$ ; thus, the conclusion is the result of evaluation among the experts has strong relevancy, (2) based on criteria validity, the test item is valid 56,7% (17 items), (3) the quality of test items based on reliability has low reliability coefficient 0.66%, (4) the quality of test items based on differentiator; the item which has worse differentiator = 2 items or 6.7%, poor = 4 items or 13.3%, good = 7 items or 23.3% and very good = 17 items or 56.7%, (5) the quality of items based on level of difficulty, 56,7% (17 items) is categorized as difficult, and 43,3% (13 items) is categorized as moderate, (6) the quality of the test items based on analysis of distractor effectiveness, there are 4 items of dysfunctional distracter.

**Keywords:** *quality of semester test analysis, classical theory*

## 1. Pendahuluan

Pembelajaran sebagai suatu sistem meniscayakan adanya evaluasi, karena evaluasi merupakan komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Menurut Edwin dan Gerald (Sudaryono, 2012:38) bahwa “evaluasi yaitu suatu tindakan atau proses untuk menentukan nilai dari sesuatu”. Nitko dan Brookhart (Mansyur, 2015:9) mendefinisikan “evaluasi sebagai suatu proses penetapan nilai yang berkaitan dengan kinerja dan hasil karya peserta didik. Berdasarkan beberapa pandangan yang telah diuraikan sebelumnya, penulis menyimpulkan bahwa evaluasi pembelajaran adalah tindakan sistematis dan berkelanjutan untuk menentukan kualitas dari proses pembelajaran berdasarkan pertimbangan dan kriteria dalam rangka pembuatan keputusan berikutnya.

Analisis kualitas soal merupakan kegiatan penting dalam upaya memperoleh instrumen penilaian yang berkategori baik. Melalui analisis kualitas soal dapat diidentifikasi butir-butir soal manakah yang termasuk dalam kategori baik, kurang baik, dan jelek. Analisis kualitas soal merupakan pekerjaan yang sangat penting bagi pembangunan instrumen evaluasi berupa tes, sebab akan berkorelasi terhadap seberapa besar alat tes tersebut dapat menggambarkan sampel perilaku dan menghasilkan nilai yang objektif serta akurat. Jika tes yang digunakan guru berkualitas buruk, maka hasil yang diperoleh tentunya kurang baik dan akan merugikan peserta didik itu sendiri. Menurut Sukiman (2012:157) bahwa “sebelum

digunakan, butir-butir soal hendaklah dianalisis terlebih dahulu agar memenuhi syarat sebagai alat evaluasi yang memiliki kualitas tinggi”. Menurut (Sujana, 2006), Analisis kualitas soal adalah suatu prosedur sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang telah disusun.

Menurut Kusaeri dan Suprananto (2012: 178), *Item and Test Analysis* (ITEMAN) merupakan program lunak (software) yang dibuat melalui bahasa pemrograman komputer yang diciptakan khusus untuk analisis statistik butir soal dan tes. Program ini dibuat dengan pendekatan analisis statistik secara klasikal yang berguna untuk menentukan kualitas butir soal dan tes berdasarkan data empiris dan hasil uji coba. Program ini dapat digunakan untuk: (1) menganalisis *data file* jawaban butir soal yang dihasilkan melalui manual entry data atau dari mesin *scanner*; (2) menskor dan menganalisis data soal pilihan ganda dan skala Likert untuk 30.000 siswa dan 250 butir soal; (3) menganalisis sebuah tes yang terdiri dari 10 skala (subtes) dan memberikan informasi tentang validitas setiap butir (daya pembeda, tingkat kesukaran, proporsi jawaban pada setiap *option*), reliabilitas (KR-20/Alpha), *standar error of measurement*, *mean*, *variance*, standar deviasi, *skew*, kurtosis untuk jumlah skor pada jawaban benar, skor minimum dan maksimum, skor median, dan frekuensi distribusi skor.

## 2. Metode Penelitian

Jenis penelitian ini adalah penelitian deskriptif. Dalam penelitian ini yang digambarkan adalah karakteristik Soal Ulangan Akhir Semester pada sekolah

dasar di Kecamatan Malili untuk Mata Pelajaran Matematika Kelas VI yang masih menerapkan KTSP. Karakteristik yang dimaksud dalam penelitian ini adalah validitas isi, tingkat kesukaran, daya beda, efektivitas pengecoh, validitas internal dan reliabilitas berdasarkan teori tes klasik.

Penelitian ini dilaksanakan di sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur pada empat gugus yang masih menerapkan KTSP. Pelaksanaan penelitian ini dilakukan pada bulan Januari – Februari 2019.

Subjek dalam penelitian ini adalah semua isi jawaban peserta didik pada lembar jawaban yang berjumlah 622 lembar Mata Pelajaran Matematika Kelas VI yang tersebar di 21 sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur.

Teknik yang dipergunakan dalam mengumpulkan data adalah teknik dokumentasi. Teknik ini dipakai untuk mendapatkan seperangkat tes, hasil pekerjaan peserta didik berupa lembar jawaban, dan daftar nama-nama peserta didik pada masing-masing sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur Semester Ganjil Tahun Pelajaran 2018/2019.

Selanjutnya, hasil validasi pakar dianalisis dengan menggunakan analisis validitas. Relevansi kedua pakar secara menyeluruh merupakan validitas isi Gregory (Ruslan, 2009:19), yaitu koefisien validitas isi. Koefisien validitas isi dapat dihitung menggunakan rumus:

$$\text{Validitas Isi: } \frac{D}{A+B+C+D}$$

Untuk menentukan apakah Soal Ulangan Akhir Semester Sekolah Dasar untuk Mata Pelajaran Matematika Kelas VI Kecamatan Malili Kabupaten Luwu Timur Tahun Pelajaran 2018/2019 memiliki derajat validitas yang cukup, maka digunakan model kesepakatan kedua penilai meyakini sebuah butir cukup relevan (relevansi kuat) yang akan ditempatkan pada sel D yang merefleksikan kesepakatan yang valid

antara penilai. Jika hasil dari koefisien validitas isi di atas 75%, maka dapat dinyatakan bahwa hasil pengukuran yang dilakukan adalah valid (Ruslan, 2009:19)

Analisis soal secara empirik dilakukan dengan menelaah karakteristik setiap butir atau item, berdasarkan teori tes klasik. Analisis ini dilakukan dengan menggunakan bantuan program (*software*) komputer *Test Analysis Program* (TAP).

Kriteria yang digunakan untuk mengetahui tingkat kesukaran butir menurut Mansyur, Rasyid & Suratno (2015:183) adalah sebagai berikut:

$$\begin{aligned} p \leq 0,30 &= \text{butir sukar} \\ 0,30 < p \leq 0,70 &= \text{butir sedang} \\ p > 0,70 &= \text{butir mudah} \end{aligned}$$

## 1. Daya beda

Daya pembeda sekurang-kurangnya harus berkualitas cukup. Kriteria yang digunakan untuk menentukan indeks pembeda pada penelitian ini yaitu seperti yang dikemukakan oleh Ebel (1979: 267) dalam Mansyur, dkk. (2015: 191):

$$\begin{aligned} 0,4 \leq D \leq 1,0 &= \text{Sangat Baik} \\ 0,3 \leq D \leq 0,39 &= \text{Baik} \\ 0,2 \leq D \leq 0,29 &= \text{Kurang} \\ &= \text{(perlu revisi)} \\ D < 0,2 &= \text{Tidak Baik} \end{aligned}$$

## 2. Efektivitas pengecoh

Efektivitas pengecoh ditentukan berdasarkan pola penyebaran pilihan jawaban untuk mengetahui pengecoh berfungsi atau tidak dengan mengacu pada kriteria. Menurut Kusaeri dan Suprananto (2012: 177), suatu pilihan jawaban atau pengecoh dapat dikatakan berfungsi apabila:

- pengecoh paling tidak dipilih oleh 5% peserta tes atau peserta didik
- pengecoh lebih banyak dipilih oleh kelompok peserta didik yang belum memahami materi yang diujikan.

### 3. Hasil Penelitian dan Pembahasan

#### 1. Deskripsi validitas isi tes

Indikator pertama yang ditinjau untuk mengidentifikasi kualitas suatu tes yaitu validitas isi untuk mengetahui sejauh mana kesesuaian butir-butir tes dengan standar kompetensi, kompetensi dasar, indikator dan materi yang tercantum dalam standar isi Mata Pelajaran Matematika Kelas VI sekolah dasar.

Untuk memperoleh informasi tentang validitas isi soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur tahun pelajaran 2018/2019, maka dilakukan estimasi terhadap validitas isi tes tersebut. Metode atau teknik estimasi yang digunakan dalam penelitian ini adalah metode atau teknik *Gregory*. Metode ini merupakan metode estimasi validitas isi tes dengan melalui penilaian dua orang pakar yang dianggap ahli dalam bidang pengukuran dan substansi materi Matematika.

Indeks kesepakatan pakar yang diperoleh tersebut kemudian dimaknai dengan mengkonfirmasi dengan sebuah kriteria. Kriteria yang digunakan untuk menentukan validitas isi tes dalam penelitian ini, yaitu jika koefisien validitas isi lebih besar dari 0,75 atau di atas 75%, maka tes tersebut dinyatakan valid secara isi.

Tabel 1 Hasil Kesepakatan Antar Dua Pakar Soal UAS MATEMATIKA

	Validator I	
	Tidak relevan Skor (1 – 2)	Relevan Skor (3 – 4)
Validator II	Tidak relevan Skor (1 – 2)	0 2 butir (36, 40)
	Relevan Skor (3 – 4)	37 butir (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 37, 38 dan 39)

Berdasarkan penilaian yang diberikan oleh kedua validator pada Tabel 1 dapat dihitung tingkat kesahihannya berdasarkan rumus validitas isi sebagai berikut:

$$\text{Validitas Isi} = \frac{D}{A + B + C + D} = \frac{37}{40} = 0,925$$

Berdasarkan hasil validasi dari dua orang pakar terhadap 40 butir soal Ulangan Akhir Semester Mata Pelajaran Matematika tersebut di atas diperoleh informasi bahwa koefisien validitas isi tes sebesar 0,925. Oleh karena koefisien validitas isi tersebut lebih besar dari 0,75, maka dapat diputuskan bahwa soal tersebut dinyatakan valid secara isi.

Adapun catatan-catatan dari dua validator ahli (*expert judgement*) adalah:

- Pada soal nomor 17 kalimat pada indikator soal sebaiknya dibuat lebih spesifik lagi.
- Nomor 36, konsep rata-rata yang dimaksud dalam soal ini tidak tepat.
- Nomor 40, kompetensi dasar dan indikator soal tidak relevan dan soalnya tergolong sangat mudah.

Berdasarkan hasil analisis validasi isi instrumen soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar, diperoleh informasi bahwa penilaian yang diberikan oleh kedua pakar atau ahli, menunjukkan bahwa pada butir 17, validator pertama memberikan penilaian relevansi lemah sedangkan validator kedua memberikan penilaian relevansi kuat sehingga dimasukkan ke dalam (sel C) sedangkan butir 36 dan 40, validator pertama memberikan penilaian relevansi kuat sedangkan validator kedua memberikan penilaian relevansi lemah sehingga dimasukkan kedalam (sel D). Hasil penilaian dari kedua pakar tersebut sehingga memberikan penilaian relevansi lemah adalah konsep dari materi dan perbaikan pada indikator soal. Sejalan dengan pendapat Azwar (2014: 175), mengemukakan pendapat bahwa pengujian validitas isi dilakukan untuk memeriksa apakah masing-masing aitem telah sesuai

dengan indikator yang hendak diungkapnya.

## 2. Deskripsi tingkat kesukaran butir

Aspek lain yang ditinjau sebagai indikator untuk memperoleh informasi terkait kualitas tes dalam penelitian adalah tingkat kesukaran butir berdasarkan teori tes klasik. Berdasarkan hasil analisis data diperoleh karakteristik 40 butir tes hasil respon siswa dari 622 orang ditinjau dari parameter tingkat kesukaran butir. Berikut hasil analisis tingkat kesukaran disajikan dalam Tabel 2.

Tabel 2 Deskripsi Hasil Analisis Tingkat Kesukaran Butir Berdasarkan Teori Tes Klasik.

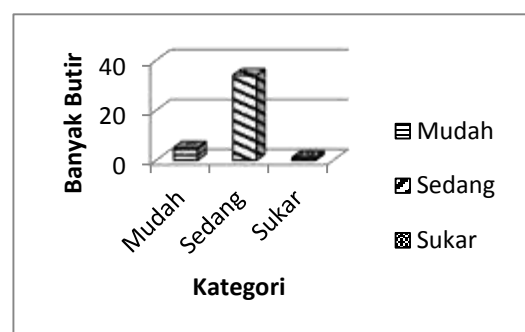
No	Kategori	Jumlah	Nomor Butir Soal	(%)
1	Mudah	5	1, 2, 5, 31 dan 40	12,5
2	Sedang	34	3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38 dan 39	85
3	Sukar	1	14	2,5
<b>Jumlah</b>		<b>40</b>		<b>100</b>

Berdasarkan Tabel 2 dapat diperoleh informasi bahwa tingkat kesukaran dari 40 butir tes, terdapat 5 butir atau 12,5% berada pada kategori mudah, 34 butir atau 85% berada pada kategori sedang, dan 1 butir atau 2,5% berada pada kategori sukar.

Hasil *out put* analisis dengan program TAP memperlihatkan bahwa butir dengan tingkat kesukaran yang paling rendah berada pada kategori mudah adalah butir 40 dengan indeks tingkat kesukaran 0,84. Hal ini menunjukkan bahwa terdapat 84% atau 521 dari 622 siswa yang menjawab dengan benar butir tersebut dan hanya 101 atau 16% siswa yang menjawab salah. Sedangkan butir dengan tingkat kesukaran yang paling tinggi berada pada kategori sukar adalah butir 14 dengan

indeks tingkat kesukaran 0,27 atau butir tersebut dapat dijawab dengan benar oleh 27% atau 165 siswa dari 622 peserta tes atau 73% siswa menjawab salah.

Apabila dilihat dari rata-rata (*mean*) indeks tingkat kesukaran butir keseluruhan diperoleh indeks kesukaran sebesar 0,507. Hal ini menunjukkan bahwa tingkat kesukaran butir berada pada kategori sedang. Hal ini juga berarti terdapat 50,7% yang dapat menjawab dengan benar semua soal ulangan akhir semester ganjil yang diberikan. Data pada Tabel 4.2 di atas dapat disajikan dalam bentuk diagram. Jika data tersebut disajikan dalam bentuk diagram maka perbandingan dari sebaran butir ditinjau dari tingkat kesukaran butir dapat dilihat pada Gambar 1 berikut ini.



Gambar 1 Diagram Pesebaran Tingka Kesukaran Butir Berdasarkan Teori Tes Klasik.

Selanjutnya ditinjau dari indikator materi, butir 1 dengan indikator siswa dapat menentukan bilangan yang belum diketahui pada operasi perkalian yang menggunakan sifat komutatif sedangkan butir 14 dengan indikator siswa dapat menentukan debit dari soal cerita yang disajikan dengan demikian dapat disimpulkan bahwa kedua butir berbeda pada taraf kemampuan yaitu butir 1 berada pada kemampuan aplikasi sedangkan pada butir 14 berada pada kemampuan analisis. Sejalan dengan pendapat Azwar (2014: 61) mengemukakan bahwa taraf kompetensi yang lebih tinggi, yang biasanya diikuti pula oleh meningkatkan taraf kesukaran

aitem. Jika ditinjau berdasarkan *mean* atau rata-rata tingkat kesukaran butir diperoleh besarnya koefisien  $p$  sebesar 0,507. Hal ini menunjukkan bahwa rata-rata tingkat kesukaran butir pada soal ulangan akhir semester berada pada kategori sedang. Senada dengan Mardapi (2017:173) tingkat kesulitan, bila menggunakan acuan norma tingkat kesulitan butir yang diterima adalah 0,30 sampai 0,80.

### 3. Deskripsi daya beda butir

Aspek kedua yang ditinjau untuk mendapatkan informasi terkait kualitas tes dalam penelitian ini adalah daya beda butir berdasarkan teori tes klasik. Berdasarkan hasil pengolahan data pada *Discrimination Index*, diperoleh karakteristik 40 butir hasil respon siswa dari 622 orang dilihat dari parameter daya beda butir disajikan dalam Tabel 3.

Tabel 3 Deskripsi Hasil Analisis Daya Beda Butir Berdasarkan Teori Tes Klasik.

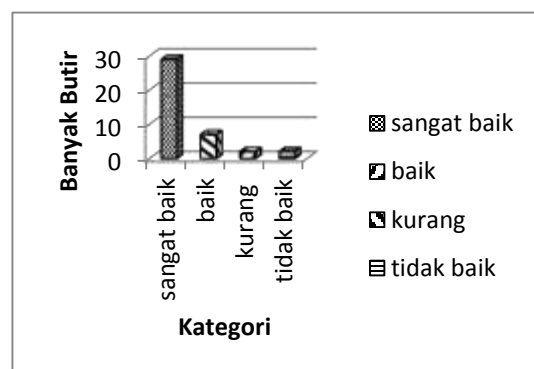
No	Kategori	Jumlah	Nomor Butir Soal	(%)
1	Sangat Baik	29	1, 2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 33, 35, 36, 37 dan 39.	72,5
2	Baik	7	14, 16, 26, 28, 32, 34 dan 38	17,5
3	Kurang	2	11 dan 40	5
4	Tidak Baik	2	5 dan 31	5
<b>Jumlah</b>		<b>40</b>		<b>100</b>

Berdasarkan Tabel 3 dapat diperoleh informasi bahwa daya beda butir, terdapat 29 butir atau 72,5% berada pada kategori sangat baik, 7 butir atau 17,5% berada pada kategori baik atau butir-butir tersebut dapat diterima tanpa revisi, 2 butir atau 5% berada pada kategori kurang atau butir-butir tersebut dapat diterima dengan

revisi dan 2 butir atau 5% berada pada kategori tidak baik.

Butir-butir yang memiliki indeks daya beda butir atau *discrimination index* ( $D$ ) yang sangat rendah, yaitu butir 5 dan 31 dengan indeks daya beda sebesar 0,10 dan 0,15. Menunjukkan bahwa butir tersebut tidak baik (buruk/tidak berfungsi). Hal ini menunjukkan bahwa butir tersebut tidak baik dalam hal membedakan antara peserta tes berkemampuan tinggi dengan peserta tes berkemampuan rendah. Jika ditinjau dari indikator soal, kompetensi atau kemampuan yang hendak diukur pada butir 5 yaitu materi bilangan kubik dengan kompetensi yang diukur adalah Menentukan akar pangkat tiga suatu bilangan kubik. Pada butir nomor 31, dengan kompetensi yang ingin diukur adalah Mengumpulkan dan membaca data.

Untuk melihat lebih jelas perbandingan persebaran dari 40 butir soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur tahun pelajaran 2018/2019 tersebut berdasarkan daya beda butir menurut teori tes klasik, maka data pada Tabel 3 di atas dapat disajikan dalam bentuk diagram. Jika data tersebut disajikan dalam bentuk diagram maka perbandingan dari sebaran butir ditinjau dari daya bedanya dapat dilihat pada Gambar 2.



Gambar 4.2 Diagram Persebaran Daya Beda Butir Berdasarkan Teori Tes Klasik

Berdasarkan hasil analisis daya beda butir pada soal ulangan akhir semester diperoleh bahwa dari 40 butir terdapat 29 butir dalam kategori sangat baik, 7 butir dalam kategori baik, 2 butir dengan kategori kurang namun bisa diterima dengan revisi dan 2 butir dalam kategori tidak baik yang sebaiknya tidak dipertahankan. Kedua butir dengan kategori tidak baik tersebut masing-masing dengan cakupan dan subtansi materi yang berbeda. Butir 5 dan 31 masing-masing dengan nilai koefisien 0,10 dan 0,15 termasuk dalam kategori tidak baik. Kedua butir tersebut semakin mendekati indeks diskriminasi 0 yang menunjukkan bahwa fungsi item tersebut tidak jelas dalam hal membedakan mana subjek yang menguasai bahan pelajaran dan mana subjek yang belum menguasai bahan pelajaran. Jika ditinjau dari materi bahwa butir nomor 5 dengan cakupan materi menyelesaikan masalah yang melibatkan operasi hitung termasuk penggunaan akar dan pangkat. Pada butir tersebut kemampuan yang hendak diukur adalah dapat menentukan arti dari bilangan pangkat tiga. Untuk menentukan arti bilangan pangkat tiga diperlukan beberapa kemampuan dan pengetahuan prasyarat dari peserta didik. Diperlukan kemampuan operasi hitung bilangan bulat, sifat-sifat operasi bilangan bulat, operasi hitung campuran serta bilangan kubik. Sedangkan pada butir nomor 31 dengan cakupan materi berkaitan menafsirkan sajian data. Pada butir tersebut kemampuan yang hendak diukur adalah berdasarkan data tentang jenis makanan dalam bentuk tabel, siswa dapat menentukan jenis makanan yang paling disukai. Oleh karena pada butir 31, diperlukan beberapa pengetahuan prasyarat dari peserta didik yaitu penyajian data (tabel, batang, lingkaran dan gambar) dan penghitungan data (nilai rata-rata, modus dan urutan data).

Berdasarkan data hasil analisis seperti yang yang dikemukakan di atas, sebanyak 29 butir dalam kategori yang sangat baik dan 7 butir dengan kategori

baik tersebut berada pada indeks diskriminasi dengan nilai  $D$  lebih dari 0,30 sedangkan 2 butir dengan kategori kurang namun bisa diterima dengan revisi, berada pada nilai  $D$  kurang dari 0,30.

Kisaran indeks diskriminasi adalah antara 1,00 dan -1,00 kemungkinan alasan untuk diskriminasi negatif saya memasukkan kunci jawaban yang salah lebih dari satu jawaban yang benar atau pertanyaan yang tidak jelas.

#### 4. Deskripsi efektivitas pengecoh

Berikut hasil analisis efektifitas pengecoh ditinjau berdasarkan frekuensi dan persentase respon yang diberikan oleh peserta tes secara keseluruhan (*total item frequencies and percentages*), hasilnya tersebut dapat dilihat pada Tabel 4 berikut:

Tabel.4 Deskripsi Hasil Analisis Efektifitas Pengecoh Butir Berdasarkan Total Frekuensi dan Persentase Respon Peserta Tes dengan Teori Tes Klasik

No	Kategori	Jumlah	Nomor Butir Soal	(%)
1	Berfungsi efektif	37	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38 dan 39.	92,5
2	Tidak berfungsi efektif	3	1, 5 dan 40	7,5
<b>Jumlah</b>		<b>40</b>		<b>100</b>

Berdasarkan Tabel 4 dapat diperoleh informasi bahwa dari 40 butir diperoleh informasi terkait efektivitas pengecoh yang ditinjau berdasarkan frekuensi dan persentase secara keseluruhan respon yang diberikan oleh peserta tes terdapat 37 butir atau sekitar 92,5 % butir yang memiliki pengecoh berfungsi efektif dan 3 butir atau 7,5 % butir yang memiliki pengecoh yang tidak

efektif. Butir yang memiliki pengecoh yang tidak efektif yaitu butir nomor 1, 5 dan 40. Pada butir tersebut kemampuan peserta didik yang hendak diukur adalah menentukan bilangan yang belum diketahui pada operasi perkalian yang menggunakan sifat komutatif, menentukan arti dari bilangan pangkat tiga dan menentukan jumlah anak yang gemar pencak silat dan karate dari data jenis cabang olahraga yang digemari.

Sedangkan jika ditinjau berdasarkan *item frequencies and percentages* respon masing-masing *group* atau kelompok, yakni kelompok berkemampuan rendah (*low group*) dan kelompok berkemampuan tinggi (*high group*) diperoleh informasi tentang efektivitas pengecoh butir seperti yang disajikan pada Tabel 5.

Tabel 5 Deskripsi Hasil Analisis Efektivitas Pengecoh Butir Berdasarkan Frekuensi dan Persentase Masing-Masing Kelompok dengan Teori Tes Klasik

No	Kategori	Jumlah	Nomor Butir Soal	(%)
1	Berfungsi dengan baik	39	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 dan 40	97,5
2	Berfungsi kurang baik/buruk	1	14	2,5
<b>Jumlah</b>		<b>40</b>		<b>100</b>

Berdasarkan Tabel 5 dapat diperoleh informasi bahwa dari 40 butir diperoleh informasi terkait efektivitas pengecoh yang ditinjau berdasarkan frekuensi dan presentase respon peserta tes masing-masing kelompok (*group*)

berkemampuan tinggi dan kelompok berkemampuan rendah, terdapat 39 butir atau 97% butir yang memiliki efektivitas pengecoh berfungsi dengan baik dan 1 butir atau 2,5% butir yang memiliki efektivitas pengecoh berfungsi tidak baik.

Untuk mengetahui tingkat kualitas tes ditinjau berdasarkan parameter efektivitas pengecoh masing-masing butir, baik berdasarkan frekuensi dan persentase respon peserta tes secara keseluruhan maupun berdasarkan frekuensi dan persentase respon peserta tes masing-masing kelompok (*group*) dijelaskan pada uraian berikut.

Untuk butir 5 nampaknya tidak semua pengecoh berfungsi dengan baik. Hal tersebut dapat dilihat pada besarnya frekuensi dan persentase secara keseluruhan respon peserta tes untuk masing-masing pengecoh. Pengecoh A dipilih oleh 66 (0,106) peserta tes dan pengecoh D dipilih 87 (0,140) peserta tes. Dengan demikian pengecoh A dan D berfungsi efektif karena hanya dipilih oleh peserta di atas 0,05. Hanya pengecoh C yang tidak berfungsi efektif yang dipilih 26 (0,042) peserta tes. Dengan demikian pengecoh C tidak berfungsi efektif karena dipilih oleh peserta tes dibawah 0,05.

Sedangkan Untuk butir 14 tampak bahwa semua pengecoh memiliki efektivitas yang cukup tinggi. Informasi tersebut diperoleh dari frekuensi dan persentase secara keseluruhan respon peserta tes semua pengecoh memiliki efektivitas di atas 5%. Pengecoh A, C dan D masing-masing memiliki frekuensi dan presentase 150 (0,241), 175(0,281) dan 132(0,212). Apabila ditinjau dari frekuensi dan presentase respon masing-masing kelompok berkemampuan rendah dan kelompok berkemampuan tinggi diperoleh informasi bahwa hanya pengecoh opsi D berfungsi kurang baik karena lebih banyak dipilih oleh kelompok berkemampuan tinggi atau kelompok atas.

Berdasarkan hasil analisis efektivitas pengecoh pada soal ulangan akhir semester ganjil Mata Pelajaran



Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur Tahun Pelajaran 2018/2019 diperoleh bahwa sebagian besar efektivitas pengecoh di atas 0,05 atau 5%. Sehingga dapat disimpulkan bahwa sebagian besar efektivitas pengecoh berfungsi dengan baik. Dari 40 butir soal hanya terdapat 2 butir yang memiliki pengecoh tidak berfungsi yaitu butir 1 dan 5. Butir 1 pada pilihan jawaban A hanya dipilih 20 orang atau 3% dan pada pilihan B hanya dipilih 27 orang atau 4%. Sedangkan pada butir 5, pada pilihan C hanya dipilih 26 orang atau 4%. Substansi materi yang diukur pada butir 1 dan 5 adalah menentukan bilangan yang belum diketahui pada operasi perkalian yang menggunakan sifat komutatif dan dapat menentukan arti dari bilangan pangkat tiga.

Adapun pengecoh 38 butir yang lain berfungsi sebagai faktor tebakkan. Efektivitas distraktor atau faktor tebakkan yang berfungsi adalah butir yang pengecohnya dipilih oleh peserta tes minimal 5%. Hal ini sesuai dengan pendapat Mardapi (2017:111), mengemukakan bahwa distribusi respons jawaban, diterima bila tiap option ada yang menjawab paling sedikit 5% dari peserta tes. Berdasarkan hasil penelitian diperoleh bahwa terdapat satu butir yang tidak berfungsi sebagai pengecoh yaitu butir 14. Butir 14 pada pengecoh D (bukan merupakan kunci jawaban) telah dipilih oleh lebih banyak kelompok tinggi daripada kelompok rendah.

#### **4. Pembahasan**

##### **a. Kualitas tes buatan guru mata pelajaran matematika kelas V SD ditinjau dari Validitas Isi**

Pada proses validasi dua pakar menunjukkan bahwa hasil penilaian terhadap 30 butir tes buatan guru mata pelajaran Matematika kelas V Sekolah Dasar, terdapat 86,67% butir telah memenuhi validitas isi. Aspek yang tidak terpenuhi untuk butir nomor 11 dan 12

yakni aspek konstruksi pada point ke 7 tentang pilihan jawaban yang berbentuk angka harus disusun berdasarkan urutan besar kecilnya nilai angka tersebut yang disusun secara kronologis. Pada butir nomor 16 dan 18, yakni aspek konstruksi pada point 1 yaitu pokok butir tidak dirumuskan secara jelas dan tegas. Proses validitas dua pakar menunjukkan bahwa hasil penilaian dari kedua validator memiliki “relevansi kuat” dengan koefisien validasi isi 0,87 atau  $V_i = 0,87 > 0,75$ . Namun ada beberapa butir soal yang belum memenuhi beberapa aspek perlu direvisi dalam hal penulisannya, Sehingga validitas soal tes buatan guru tersebut dapat disimpulkan tinggi.

##### **b. Kualitas tes buatan guru mata pelajaran matematika kelas V SD ditinjau dari Validitas Kriteria**

Berdasarkan hasil analisis Soal Ujian Semester Ganjil Mata Pelajaran Matematika kelas V Sekolah Dasar terdapat 56,7% dinyatakan valid. Berdasarkan uraian tersebut dapat disimpulkan bahwa Tes Buatan Guru Mata Pelajaran Matematika kelas V SD Negeri di Kecamatan Wotu Tahun Pelajaran 2018/2019 memiliki kualitas yang baik dari segi validitasnya. Butir yang valid dapat disimpan di bank soal. Sebaliknya, butir yang tidak valid perlu diperbaiki dengan menyesuaikan indikator dan meningkatkan penguasaan teknik tentang penyusunan butir.

##### **c. Kualitas tes buatan guru mata pelajaran matematika kelas V SD Timur ditinjau dari Reliabilitas**

Berdasarkan perhitungan yang dilakukan secara manual diperoleh hasil  $r_{11} = 0,667$ . Hasil perhitungan tersebut menunjukkan bahwa soal Tes Buatan Guru Mata Pelajaran Matematika V Sekolah Dasar memiliki tingkat reliabilitas yang rendah karena memiliki koefisien reliabilitas ( $r_{11}$ )  $< 0,70$ . Rendahnya koefisien reliabilitas soal karena terbatasnya jumlah butir yang dibuat

oleh guru, sehingga guru dapat menambahkan jumlah butir yang valid.

Berdasarkan uraian di atas disimpulkan bahwa Tes Buatan Guru Mata Pelajaran Matematika kelas V SD Negeri di Kecamatan Wotu Tahun Pelajaran 2018/2019 merupakan soal yang memiliki tingkat reliabilitas yang rendah.

**d. Kualitas tes buatan guru mata pelajaran matematika kelas V SD ditinjau dari Tingkat Kesukaran**

Berdasarkan hasil analisis diperoleh butir dengan persentase 56,7% yang termasuk sukar, 43,3% tergolong dalam butir sedang, Butir yang sukar akan membuat siswa putus asa untuk mencoba lagi karena diluar kemampuan siswa. Suatu butir dianggap baik jika memiliki tingkat kesukaran antara 0,30 – 0,70 (Arifin. 2013: 225). Sehingga dapat disimpulkan bahwa soal Tes Buatan Guru Mata Pelajaran Matematika kelas V Sekolah Dasar termasuk soal yang memiliki tingkat kesukaran yang tinggi dengan presentase 56,7%.

**e. Kualitas tes buatan guru mata pelajaran matematika kelas V SD ditinjau dari Daya Pembeda**

Berdasarkan hasil analisis, terdapat 56,7% memiliki daya pembeda sangat baik, 23,3% memiliki daya pembeda yang baik, 13,3 % memiliki daya pembeda kurang baik ( perlu revisi) dan 6,7% memiliki daya pembeda buruk/jelek.

Berdasarkan pembahasan di atas, disimpulkan bahwa Tes Buatan Guru Mata Pelajaran Matematika kelas V SD Negeri di Kecamatan Wotu Tahun Pelajaran 2018/2019 memiliki daya pembeda yang sangat bagus. Hanya ada 6 butir atau sebesar 20% memiliki daya pembeda yang buruk/jelek. Kemudian ada 24 butir atau sebesar 80% yang memiliki daya pembeda kurang baik, baik, dan sangat baik.

**f. Kualitas tes buatan guru mata pelajaran matematika kelas V SD ditinjau dari Efektivitas Pengecoh**

Berdasarkan hasil analisis data diperoleh 26 butir soal yang berfungsi baik dan 4 butir soal yang berfungsi kurang baik. Butir yang kurang berfungsi dalam efektivitas pengecoh adalah butir 5, 21, 27 dan 29.

**5. Kesimpulan**

Berdasarkan hasil penelitian dan pembahasan yang telah dikemukakan dalam penelitian ini, dapat disimpulkan bahwa:

1. Validitas isi soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur tahun pelajaran 2018/2019 telah valid secara isi.
2. Tingkat kesukaran butir soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur tahun pelajaran 2018/2019 berada pada kategori sedang dengan proporsi jumlah soal kategori mudah 5 butir (12,5%), kategori sedang 34 (85% ) dan kategori sukar 1 butir (2,5%).
3. Daya beda butir pada soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur Tahun Pelajaran 2018/2019 diperoleh 29 butir dalam kategori sangat baik, 7 butir dalam kategori baik, 2 butir kategori kurang dan 2 butir dalam kategori tidak baik.
4. Efektivitas pengecoh pada soal ulangan akhir semester ganjil Mata Pelajaran Matematika Kelas VI sekolah dasar di Kecamatan Malili Kabupaten Luwu Timur Tahun Pelajaran 2018/2019, terdapat 37 butir yang pengecohnya berfungsi dan 3 butir yang pengecohnya tidak berfungsi.

## 6. Daftar Pustaka

- Ali, S. & Khaeruddin. 2012. *Evaluasi Pembelajaran*. Makassar: Badan Penerbit UNM.
- Anastasi, A. & Urbina, S. 2007. *Tes Psikologi*. Terjemahan oleh Hariono, R. S. I. Jakarta: PT. Indeks.
- Arifin, Z. 2017. *Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Azwar, S. 2014. *Tes Prestasi Fungsi dan Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: Pustaka Pelajar.
- BSNP. 2007. *Standar Kompetensi dan Kompetensi Dasar Tingkat SD/MI Kelas VI Mata Pelajaran Matematika*. Jakarta: Dirjen Mendikdasmen Direktorat Pembinaan TK dan SD.
- Kusaeri & Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Mansyur, Rasyid, H., & Suratno. 2015. *Assesmen Pembelajaran di Sekolah*. Yogyakarta: Pustaka Pelajar.
- Mardapi, D. 2017. *Pengukuran Penilaian & Evaluasi Pendidikan*. Yogyakarta: Parama Publishing.
- Ruslan. 2009. Validitas Isi. *Buletin Pa'biritta*. No. 10. Tahun VI, 18 - 19.
- Sudaryono. 2012. *Dasar-Dasar Evaluasi Pembelajaran*. Yogyakarta: Graha Ilmu
- Sujana. 2006. *Penilaian Hasil Belajar Proses Mengajar*. Bandung: PT Remaja Rosdakarya
- Sukardi. 2008. *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Yogyakarta: Bumi Aksara.
- Sukiman. 2012. *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Madani.